



European
Commission

JRC SCIENCE AND POLICY REPORTS

Environmental Performance Index 2014 JRC Analysis and Recommendations



Stergios Athanasoglou

Dorota Weziak-Bialowolska

Michaela Saisana

2014

Report EUR 26623 EN

Joint
Research
Centre

European Commission
Joint Research Centre
Institute for the Protection and the Security of the Citizen

Contact information
Stergios Athanasoglou, Dorota Weziak-Bialowolska, Michaela Saisana

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy
E-mail: stergios.athanasoglou@jrc.ec.europa.eu, dorota.bialowolska@jrc.ec.europa.eu,
michaela.saisana@jrc.ec.europa.eu
Tel.: +(39) 0332 786590
Fax: +(39) 0332 785733

<http://ipsc.jrc.ec.europa.eu/>
<http://www.jrc.ec.europa.eu/>

Composite Indicators website : <http://composite-indicators.jrc.ec.europa.eu>

This publication is a Science and Policy Report by the Joint Research Centre of the European Commission.

Legal Notice

This publication is a Science and Policy Report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

JRC89939

EUR 26623 EN

ISBN 978-92-79-37915-4 (PDF)

ISSN 1831-9424 (online)

doi: 10.2788/64170

Luxembourg: Publications Office of the European Union, 2014

© European Union, 2014

Reproduction is authorised provided the source is acknowledged.

Printed in Italy

Environmental Performance Index 2014

JRC Analysis and Recommendations

Stergios Athanasoglou, Dorota Weziak-Bialowolska & Michaela Saisana

European Commission, Joint Research Centre (Ispra, Italy)

Executive Summary

The latest edition of the Environmental Performance Index (EPI) was presented and discussed on January 25, 2014 during the World Economic Forum Annual Meeting in Davos (Switzerland). The EPI is released biannually since 2006 by the Yale Center for Environmental Law & Policy (YCELP) and the Center for International Earth Science Information Network (CIESIN) at Columbia University, in collaboration with the Samuel Family Foundation and the World Economic Forum.

The EPI ranks how well countries perform on high-priority environmental issues in two broad policy areas: protection of human health from environmental harm and protection of ecosystems. Within these two policy objectives the EPI benchmarks country

performance in nine issue areas comprised of 20 indicators. The selected indicators measure how close 178 countries are to meeting internationally established targets or, in the absence of agreed-upon targets, how they compare to the range of observed countries.

The Econometrics and Applied Statistics Unit at the European Commission Joint Research Centre (JRC) in Ispra, Italy, was invited for a fifth consecutive time to audit the EPI because of the changes made to the selected list of indicators and the inclusion of 46 new countries.

The JRC assessment of the 2014 EPI focused on two main questions:

- 1) Is the EPI multi-level structure statistically coherent?

- 2) What is the impact of modelling assumptions on the 2014 EPI ranking?

This report answers these questions and highlights issues where the developing team may further reflect upon.

First, our analysis showed that the 2014 EPI is well-balanced with respect to the two policy objectives and that these objectives are also adequately correlated to justify their aggregation into an overall index. This is a noteworthy improvement compared to past releases. Second, we observed good to strong correlations between indicators and the respective EPI issue areas, which implies meaningful indicator contribution to the variance of the aggregate scores. Third, the indicators' correlation structure within and across the nine EPI issue areas suggests that the indicators have been allocated to the most relevant environmental issue.

Points that call for a refinement of the 2014 EPI framework were also spotted. This refinement regards mainly three of the nine EPI issue areas – Forests, Fisheries and Agriculture – that belong to the Ecosystem Vitality objective. Although present in the conceptual framework, these three EPI issue areas do

not seem to have considerable impact on the variation of country scores, hence they do not contribute significantly to the EPI ranking. Hence, a more thorough look into the indicators under Forests, Fisheries and Agriculture is suggested.

The uncertainty analysis of the 2014 EPI ranking was based on a combination of a Monte Carlo experiment and a multi-modelling approach, following good practices suggested in the composite indicators literature. We investigated the robustness of EPI country ranks to two key choices: the policy objectives weights and aggregation function. We simulated 500 weight-aggregation pairs that ensured balanced policy objective contributions. The aggregation functions we considered belong to the family of generalized weighted means – including arithmetic and geometric means – that allow for limited substitutability between the EPI objectives on protection of human health from environmental harm and protection of ecosystems. We found that all published 2014 EPI ranks lay within the simulated 95% confidence intervals. Nevertheless, several country ranks vary significantly with changes in the policy objectives' weights and the aggregation function: 38 countries have 95%

confidence interval widths between 20 and 29, 20 countries between 30 and 39, 6 countries between 40 and 49 (India, Solomon Islands, Pakistan, Swaziland, CAR, China, Vanuatu), and 2 between 50 and 59 (Belize, India). For those countries, the 2014 EPI ranks need to be treated with caution.

The choice of aggregation function at the policy objectives level was the main driver of the variation in country ranks. Choosing the average absolute shift in rank as our robustness metric, we found that the aggregation function choice accounts for 94% of the sample variance, whilst the objectives' weights choice only for 4%. This result suggests that should the methodological choices behind the 2014 EPI stimulate further discussions, then these should focus more on the

aggregation formula for the two policy objectives and much less on their weights. However, it is also worth noting that the confidence intervals of 9 countries (Kiribati, South Africa, Argentina, Guatemala, Barbados, Uzbekistan, Libya, Zambia, and Grenada) became wider when considering the simulation results under fixed arithmetic aggregation. This suggests that the weight uncertainty for the two objectives plays a role for some countries, though not for the majority of them.

The auditing conducted herein has shown the potential of the 2014 Environmental Performance Index, upon some further refinements, in reliably identifying weaknesses and ultimately monitoring national performance in high-priority environmental issues around the world.

Table of Contents

Executive Summary	3
1. Introduction.....	7
2. Is the 2014 EPI structure statistically coherent?	12
2.1. <i>Data quality and availability</i>	<i>12</i>
2.2. <i>Statistical dimensionality and grouping of components</i>	<i>12</i>
2.3. <i>Variance-based sensitivity analysis</i>	<i>17</i>
3. How do modelling assumptions influence the 2014 EPI ranking?	19
3.1. <i>Weight uncertainty</i>	<i>20</i>
3.2. <i>Aggregation function uncertainty</i>	<i>20</i>
3.3. <i>Generating weight-aggregation samples.....</i>	<i>21</i>
3.4. <i>Uncertainty Analysis Results</i>	<i>22</i>
3.5. <i>The relative importance of uncertain input factors to the variation in 2014 EPI ranks.....</i>	<i>25</i>
3.6. <i>Uncertainty analysis under fixed arithmetic aggregation.....</i>	<i>27</i>
4. Conclusions	30
References	32

1. Introduction

The 2014 Environmental Performance Index 2014 (EPI) is a joint project between the Yale Center for Environmental Law & Policy (YCELP) and the Center for International Earth Science Information Network (CIESIN) at Columbia University, in collaboration with the Samuel Family Foundation and the World Economic Forum. The 2014 EPI is constructed through 20 indicators reflecting national-level environmental data. The EPI framework comprises two overarching policy objectives – protection of human health from environmental harm (Environmental Health) and protection of ecosystems (Ecosystem Vitality). These two policy objectives are made of three and six issue areas, respectively. The 2014 EPI framework with its three level structure – from indicators to issues areas, from issues areas to policy objectives, and from policy objectives to an overall index – is shown in Figure 1.

Data were collected for 232 countries but, due to missing or incomplete data, the 2014 EPI was calculated for 178 countries, including 46 more countries with respect to the previous release (2012 EPI). The preliminary treatment of the raw data, performed by the EPI developing team, started by:

- dividing the raw data by population, land area, gross domestic product, or other denominator to make data comparable across countries;
- transforming particularly asymmetric (skewed) indicators to better differentiate performance among countries;
- normalizing all indicators with a proximity-to-target method (0: lowest to 100: target or beyond) to make data comparable across indicators.

A country's policy category scores were calculated as the simple (unweighted) arithmetic average of the underlying normalized (transformed) indicators. Conversely, a weighted arithmetic average was used at higher aggregation levels. The weighting scheme adopted by the developing team is shown in Figure 2.



Figure 1. The 2014 EPI Framework includes 9 issues and 20 indicators. Access to Electricity is not included in the figure because it is not used to calculate country scores.

Figure 1. Framework of the EPI-2104

Source: Hsu et al. 2014, p.18

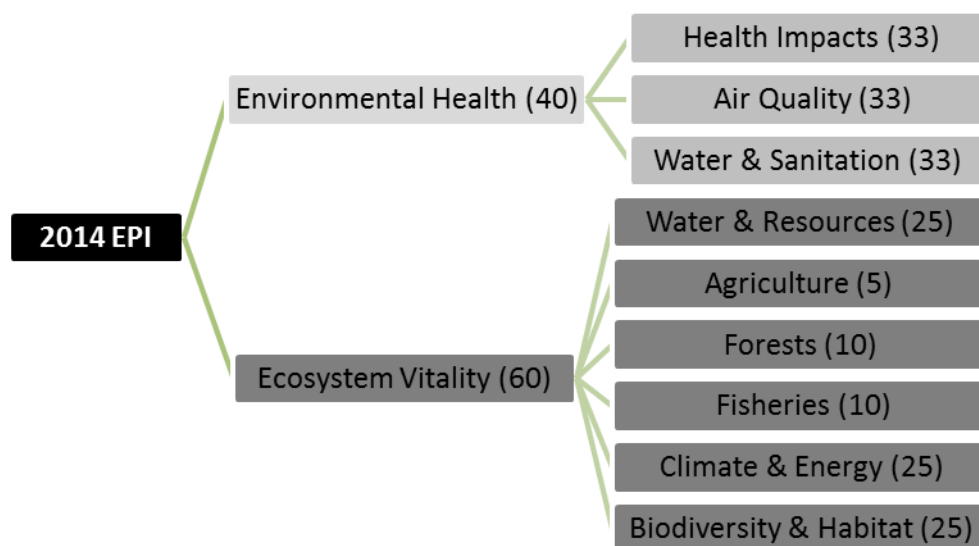


Figure 2. Weighting scheme (%) of the 2014 EPI components

The Joint Research Centre (JRC) audit starts from this point onwards, namely from the normalized indicators.¹

Inevitably, the construction of the 2014 EPI raises both conceptual and practical challenges. The conceptual challenges have been discussed in the main EPI report (Hsu et al., 2014). Herein, the focus is on the practical challenges related to the data quality and the methodological choices on the grouping of these data into nine issues areas, two policy objectives and an overall index. We consider statistical soundness to be a necessary but not a sufficient condition for a sound EPI. Given that a statistical audit of an index is mostly

¹ The Unit of Econometrics and Applied Statistics at the JRC has developed an in-house quality control process that involves both conceptual and methodological tests for the suitability and reliability of composite indicators and the development and presentation of dashboards of indicators. The JRC has helped over 60 international organizations to fine-tune their indices, such as Transparency International, Harvard, M.I.T., INSEAD, World Intellectual Property Organization, United Nations, World Economic Forum, European Central Bank, among others. Further information: <http://composite-indicators.jrc.ec.europa.eu>

based on correlations, but not only, the correspondence of a composite indicator with a real world phenomenon needs to be critically addressed whereas “correlations need not necessarily represent the real influence of the individual indicators on the phenomenon being measured” (OECD, 2008, p. 26). The point we are making here is that the validity of the EPI relies on the interplay between both statistical and conceptual soundness. In conclusion, a sound composite indicator involves an iterative process that goes back and forth between the theoretical understanding of a phenomenon on the one hand, and the empirical observations on the other.

We look into the statistical properties of the 2014 EPI by analyzing the following issues:

- 1) **Is the EPI structure statistically coherent?** Answering this question implies looking for EPI components that do not contribute significantly to the variance of their aggregates, testing whether a single measure is enough to summarize the components that are conceptually grouped in the same issue area or policy objective, analyzing whether the two EPI objectives capture different yet related aspects of a country’s performance.
- 2) **What is the impact of modelling assumptions on the 2014 EPI ranking?** This type of analysis helps to reveal for which countries the EPI rank can be taken at face value and which country ranks are more sensitive to changes in the policy objective weights and the choice of aggregation function (hence to a lower degree of substitutability) among policy categories. It also helps to identify those sources of uncertainty that are most influential in the development of the EPI and therefore to help focus discussions on those uncertainties that have an impact on the EPI results.

Regarding the first issue, the statistical coherence in the EPI framework was analyzed along five points:

1. assessment of the amount of missing values;
2. detection of indicators with strong collinearity;
3. detection of indicators that behave as noise;
4. detection of indicators that point to the opposite direction;
5. assessment of the statistical dimensionality and reliability of the components.

Point (1) was dealt with descriptive statistics. Points (2) – (4) were addressed by correlation analysis of the dataset. Finally, point (5) was assessed employing principal component analysis.

On the second part of the analysis, we investigated the robustness of 2014 EPI country ranks to the choice of policy objective weights and aggregation function via an uncertainty and sensitivity analysis. In doing so, we simulated 500 weight-aggregation pairs that ensure balanced policy objective contributions. The aggregation functions we considered belonged in the family of generalized weighted means (Decancq and Lugo, 2013), which allow for limited substitutability between the different dimensions of the index (including the arithmetic and geometric means). We found that country ranks vary significantly with such changes in weights and aggregation function: 38 countries have 95% confidence interval widths between 20 and 29, 20 countries between 30 and 39, six countries between 40 and 49 (India, Solomon Islands, Pakistan, Swaziland, CAR, China, Vanuatu), and two countries between 50 and 59 (Belize, India).

The choice of aggregation function is the main driver of this variation. Choosing the average absolute shift in rank as our robustness metric, we estimated that the choice of aggregation function is responsible for 94% of the sample variance, whilst the choice of weights account for only 4%. However, it is also worth noting that the confidence intervals of nine countries (Kiribati, South Africa, Argentina, Guatemala, Barbados, Uzbekistan, Libya, Zambia, and Grenada) become wider under fixed arithmetic aggregation. This suggests that weight uncertainty does play a role for some countries, albeit less so than the uncertainty over the choice of aggregation function, in the observed variance of the 2014 EPI ranks.

2. Is the 2014 EPI structure statistically coherent?

2.1. Data quality and availability

A complete dataset was sent to the JRC for the audit. Missing data had already been estimated by the developing team and the remaining missing cells were related to non-applicable cases and were due to country differences in natural resource endowments, physical characteristics and geographical factors. Examples include landlocked countries, for whom fisheries or marine sustainability are irrelevant, or desert countries with little to no forest cover. All non-applicable cases were found under the Ecosystem Vitality objective (within Forests, Fisheries, Climate & Energy, Biodiversity & Habitat, and Agriculture).

In the computation of the Ecosystem Vitality objective, the non-applicable cases were accounted for in the following way. For the countries with irrelevant indicators/categories, all other relevant categories received proportionally greater weight in the computation of the objective in order to preserve the same relative importance of the available categories across all countries.

The statistical features of the 2014 EPI were explored through univariate and multivariate analyses. Univariate analysis was carried out at the indicator level and focused on the presence of missing data. The data used in the JRC analysis were provided in the [0, 100] scale after correcting for highly asymmetric distributions. Nevertheless, the indicator Average Exposure to PM2.5 remains asymmetric and almost 40% of the 0-100 scale is empty². This is not necessarily a problem per se in the index construction but it is worthy commenting on as it might evidence an issue that went unnoticed.

2.2. Statistical dimensionality and grouping of components

The analysis based on correlation measures led us to the following conclusions. First, a good positive correlation between the two policy objectives ($r=.60$) and between a policy objective

² Average Exposure to PM2.5 has a skewness = -2.7 and kurtosis = 11.9. This asymmetry is driven by five scores: 1 value at 2.44 points (China) and 4 values close to 28-33 points (Nepal, Pakistan, India, Bangladesh), whilst all remaining countries get 50 points or more.

and the 2014 EPI ($r=.90$ and $r=.87$, respectively for the Environmental Health and Ecosystem Vitality) was found. This suggests that it is reasonable to further aggregate the policy objectives into an index, given that they share some common variance. The good correlation between the policy objectives in the 2014 EPI framework is a notable improvement compared to past versions of the index, where the two objectives had resulted as either not correlated at all, or even moderately negatively correlated (in those cases further aggregation into an index was not advisable, see Saisana and Saltelli, 2010 for more details). Second, we found considerably strong correlations between indicators and their corresponding category, suggesting that the indicators provide meaningful information on the variation of the category scores. However, correlations are not equally strong across indicators, thus pointing to their unequal importance within categories. This point is discussed in more detail at the next section. Third, a cross-correlation analysis within and across the nine 2014 EPI issue areas confirmed the expectation that the indicators are more correlated with their own category than with any other category. Hence, no re-allocation of the indicators to other categories is needed.

Hence, the 2014 EPI framework has notably improved compared to past versions. Nevertheless, a few challenging issues remain, which are the following:

- Within the Environmental Health objective, the Air Quality issue area is the only one that needs refinement. Although all three underlying indicators correlate considerably and with similar strength with the issue area ($r = .62-.73$), two of them, Average Exposure to PM2.5 and PM2.5 Exceedance, explain less than 1% of the variation in either the Environmental Health objective or the EPI³; hence only the Household Air Quality provides substantial information to the Environmental Health objective and the EPI ($r = .88$ and $.75$, respectively).
- Five of the six issue areas under the Ecosystem Vitality objective need further reflection and refinement.
 - a) Climate & Energy: Whilst Change of Trend in Carbon Intensity can explain some of the variance of the Climate & Energy scores ($r=.58$), the explanatory

³ The amount of variance explained by an indicator is equal to the squared value of the Pearson correlation coefficient between the indicator and the EPI component.

power of this indicator is lost at the next two levels of aggregation, namely the Ecosystem Vitality or the EPI.

- b) Biodiversity & Habitat: Two very strongly correlated indicators – National Biome Weights and Global Biome Weights ($r=0.97$)– lead to a double counting problem within the issue area. In fact, either of these two indicators accounts for 86% of the variation of the Biodiversity & Habitat scores ($r=0.93$, see Table 1), whilst the other two indicators – Marine Protected Areas and Critical Habitat Protection – explain much less ($r = .65$ and $.74$, respectively).
- c) Fisheries: The explanatory power of Fish Stock is much lower than that of Coastal Shelf Fishing Pressure ($r = .55$ and 0.95 , respectively).⁴ The explanatory power of Fish Stock is lost at the higher aggregation levels (Ecosystem Vitality, EPI).
- d) Forests: The single indicator included herein, Change in Forest Cover, does not provide any substantial information on the variation of either the Ecosystem Vitality objective or the EPI.
- e) Agriculture: The Agricultural Subsidies indicator is negatively correlated to the Pesticide Regulation, and negatively correlated with the Ecosystem Vitality and the EPI. This flags, either the existence of an undesirable trade-off, or the need to replace subsidies with a more suitable indicator.

In order to assess whether the above problematic correlations between some of the EPI components result, at least in part, from the arithmetic aggregation method, we re-calculated them under the assumption of a geometric aggregation. In particular, we employed geometric aggregation to aggregate (a) issue area scores into policy objectives scores, and (b) policy objective scores into the EPI. The results are presented in Table A2 in the Appendix. In general it appears that all problematic correlations (unexpectedly negative or not significant) remain problematic. This means that problems with orientations of the indicators and their influence on the EPI cannot be attributed, even in part, to the aggregation method.

⁴ Practically, 90% of the variance in the Fisheries scores can be explained by Coastal Shelf Fishing Pressure.

Consequently, most likely these problems cannot be addressed with the use of another aggregation formula. Instead, a careful revision of the 2014 EPI framework is warranted.

Finally, Principal Component Analysis (PCA) was used within each of the two policy objectives to assess to what extent the conceptual framework is confirmed by statistical approaches and to identify eventual pitfalls. This analysis showed that indeed the Environmental Health objective can be considered as uni-dimensional with one latent variable explaining 71.1% of variance in the three underlying issue areas. On the other hand, there are two latent dimensions in the six issue areas of the Ecosystem Vitality. The first one captures 26.35% of the variance and is described by Water Resources, Agriculture, Fisheries (in part), and Climate & Energy. The second dimension captures 23.56% of the variance and is described by the Forests, Fisheries (in part) and Biodiversity & Habitat. Hence, there is a notable amount of information that is lost when aggregating directly the six issues areas into a policy objective.

Table 1. Pearson correlation coefficients between 2014 EPI components

Policy objectives	Issue Areas	Indicators	correlation between indicator and issue area	correlation between indicator and objective	correlation between indicator and EPI	correlation between issue area and objective	correlation between objective and EPI
Environmental Health	Health Impacts	Child Mortality	1.00	0.93	0.85	0.93	0.88
	Air Quality	Household Air Quality	0.62	0.88	0.75	0.60	
		Average Exposure to PM2.5	0.73	0.02 ^{ns}	-0.06 ^{ns}		
		PM2.5 Exceedance	0.68	-0.08 ^{ns}	-0.14 ^{ns}		
	Water & Sanitation	Access to Drinking Water	0.95	0.90	0.84	0.95	
		Access to Sanitation	0.96	0.90	0.81		
Ecosystem Vitality	Water Resources	Wastewater Treatment	1.00	0.75	0.81	0.75	0.90
	Agriculture	Agricultural Subsidies	0.57	-0.53*	-0.59*	-0.02 ^{ns}	
		Pesticide Regulation	0.62	0.41	0.43		
	Forests	Change in Forest Cover	1.00	0.14 ^{ns}	0.22	0.14 ^{ns}	
	Fisheries	Coastal Shelf Fishing Pressure	0.95	-0.05 ^{ns}	-0.07 ^{ns}	-0.01 ^{ns}	
		Fish Stocks	0.55	0.12 ^{ns}	0.05 ^{ns}		
	Biodiversity & Habitat	National Biome Weights	0.93	0.58	0.41	0.67	
		Global Biome Weights	0.93	0.58	0.42		
		Marine Protected Areas	0.65	0.55	0.43		
		Critical Habitat Protection	0.74	0.44	0.34		
	Climate & Energy	Trend in Carbon Intensity	0.54	0.35	0.24	0.61	
		Change of Trend in Carbon Intensity	0.58	0.11 ^{ns}	-0.07 ^{ns}		
		Trend in CO2 Emissions per KWH	0.60	0.34	0.39		

Note: * indicates undesirable negative correlation; 'ns' indicates not significant correlation at 99% level

Table 2. Pearson correlation coefficients between the nine issue areas in the 2014 EPI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health Impacts (1)	1.00							
Air Quality (2)	0.38	1.00						
Water & Sanitation (3)	0.87	0.39	1.00					
Water Resources (4)	0.66	0.28	0.74	1.00				
Agriculture (5)	-0.06	-0.07	-0.09	-0.13 ^{ns}	1.00			
Forests (6)	0.23	0.12	0.28	0.12 ^{ns}	-0.03 ^{ns}	1.00		
Fisheries (7)	-0.08	0.14	-0.16	-0.19 ^{ns}	-0.16 ^{ns}	-0.04 ^{ns}	1.00	
Biodiversity & Habitat (8)	0.18	0.06	0.16	0.18 ^{ns}	-0.03 ^{ns}	-0.31*	0.05 ^{ns}	1.00
Climate & Energy (9)	0.19	-0.16	0.21	0.36	0.07 ^{ns}	-0.05 ^{ns}	-0.16 ^{ns}	0.10 ^{ns}

* indicates undesirable negative correlation; 'ns' indicates not significant correlation at 99% level

2.3. Variance-based sensitivity analysis

Our statistical analysis in the previous sections was based on the classical correlation coefficients. Here, we extend our assessment to a non-linear framework, to anticipate potentially legitimate criticism about the nonlinearity of the correlations between the EPI components. To this end, global sensitivity analysis has been employed in order to evaluate an issue area's contribution to the variance of the EPI scores. The overarching consideration made by the developing team was that the two objectives of the EPI should have equal importance in the overall index (this is why Environmental Health and Ecosystem Vitality are given weights of 0.4 and 0.6, respectively). Then, within the Environmental Health objective, there are three issue areas that are designed to be of equal importance. On the other hand, within the Ecosystem Vitality objective the importance of six issue areas is more diversified (see Figure 2). The Water Resources, Biodiversity & Habitat, and Climate & Energy are designed to have the same importance (i.e. 25%) and be more important with respect to any of the remaining three issue areas.

Our tests focused herein on identifying whether the 2014 EPI is statistically well-balanced in its objectives and in its issue areas within an objective. There are several approaches to test this, such as eliminating one issue area at a time and comparing the resulting ranking with the original ranking, or using a simple (e.g., Pearson or Spearman rank) correlation coefficient. A more appropriate measure aptly named 'main effect' (henceforth S_i) has been

applied here, also known as correlation ratio or first order sensitivity measure (Saltelli et al., 2008). In applying this measure to several case studies on composite indicators, Paruolo et al. (2013) argue that the suitability of Pearson's correlation ratio as a measure of the importance of variables in an index is four-fold: (a) it offers a precise definition of importance that is 'the expected reduction in variance of the composite indicator that would be obtained if a variable could be fixed', (b) it can be used regardless of the degree of correlation between variables, (c) it is model free, in that it can be applied also in non-linear aggregations, and finally (d) it is not invasive, in that no changes are made to the composite indicator or to the correlation structure of the indicators.

The results of our analysis appear in Table 3. Examining the S_i 's for the two EPI policy objectives, we see that they are almost perfectly balanced (0.82 for Environmental Health vs. 0.81 for Ecosystem Vitality). This suggests that the weighting scheme chosen by the developing team has indeed led to the desired outcome on the relative importance of the two policy objectives in the EPI.

Before proceeding with the presentation of the results at the issue area level, it must be noted that many countries lack scores on one or more issue areas within the Ecosystem Vitality. Specifically, 23, 25, and 27.5% of countries lack scores for the Forests, Fisheries, and Climate & Energy issue areas, respectively. Thus, when studying the relative importance of the six issue areas within the Ecosystem Vitality objective, as well as with respect to the 2014 EPI as a whole, we had smaller, though still sizeable (comprising at least 130 countries), datasets to work with.

With the above in mind, let us proceed with the analysis. At the issue area level, the EPI seems decidedly less balanced. Within the Environmental Health objective, the S_i for Air Quality is significantly smaller than that of Health Impacts and Water & Sanitation (0.42 vs. 0.87 and 0.90, respectively), despite their equal weights within the objective. Conversely, within the Ecosystem Vitality objective, we again notice an uneven pattern: Agriculture has an S_i of 0.29, while Forests and Fisheries are practically insignificant, despite the fact that they are weighted twice as much. At the same time, Water Resources, Biodiversity and Climate & Energy have relatively similar S_i values, in line with their equal weight.

Table 3. Importance measures for the two 2014 EPI objectives and nine issue areas

EPI Component	Importance Measures wrt EPI	EPI Weights	Importance measures wrt EPI objectives	Weights within objective	
	Si		Si		
Environmental Health	0.82	0.40	Environmental Health		
Ecosystem Vitality	0.81	0.60			
Health Impacts	0.79	0.13			0.87
Air Quality	0.22	0.13			0.42
Water & Sanitation	0.77	0.13			0.90
			Ecosystem Vitality		
Water Resources	0.74	0.15	0.56	0.25	
Agriculture	0.28	0.03	0.29	0.05	
Forests	0.05	0.06	0.04	0.1	
Fisheries	0.01	0.06	0.01	0.1	
Biodiversity	0.24	0.15	0.49	0.25	
Climate & Energy	0.22	0.15	0.38	0.25	

Notes: The S_i values are the kernel estimates of the Pearson correlation ratio, as in Paruolo et al., (2013). Problematic and mutually inconsistent values are highlighted in red.

Moving to the importance measures for the 2014 EPI as a whole, the statistical concerns of the within-objective analysis remain. Indeed, they are exacerbated. The S_i of Water Resources is significantly higher than that of Biodiversity and Climate & Energy, despite their equal weight. The S_i 's of the latter two issue areas are lower than that of Agriculture, even though their weights are five times higher. This points to a substantial imbalance.

3. How do modelling assumptions influence the 2014 EPI ranking?

Every country score on the EPI depends on subjective modelling choices: objective-issue area structure, selected variables, imputation or not of missing data, normalization, weights, aggregation method, among other elements. The robustness analysis performed by the JRC aimed at assessing the joint impact of such modelling choices on the rankings, and thus to

complement the 2014 EPI ranks with error estimates stemming from the unavoidable uncertainty in the choices made.

Our assessment of the 2014 EPI was based on a combination of Monte Carlo experiments and multi-modelling approached, following good practices suggested in the composite indicators literature (Saisana *et al.*, 2005; Saisana *et al.*, 2011). We focused on two key issues: the choice of objective weights and aggregation function. Undoubtedly, we could have incorporated other uncertain elements of the index to our robustness analysis (e.g., normalization scheme) but results from this type of analysis in past versions of the index suggested that the weights and the aggregation formula at the objective level are the two assumptions with the highest impact on the EPI ranking.

3.1. Weight uncertainty

As mentioned earlier, the 2014 EPI assigns a weight of 0.4 and 0.6 to the objectives of Environmental Health and Ecosystem Vitality, respectively. This is done to ensure that the index reflects equal importance between these two broad objectives, where importance is measured by their correlation to the overall index score. However, while this choice is guided by sound logic, it does not imply uniqueness. Indeed, it turns out that many other weighting schemes would maintain roughly equal importance of the two objectives in the index. In our uncertainty analysis, we allowed the weights of the Environmental Health objective to vary uniformly in the interval [0.3,0.5], and thus of Ecosystem Vitality in [0.5, 0.7], without altering the weight distributions at the category or indicator levels. For instance, if we drew a value of 0.45 for Environmental Health, this would imply weights of $0.45 \times 1/3 = 0.15$ for the three issue areas within this objective, and similar calculations would apply for the Ecosystem Vitality's total weight of 0.55.

3.2. Aggregation function uncertainty

Regarding the choice of aggregation formula, decision-theory practitioners have challenged the use of simple arithmetic averages because of their fully compensatory nature, in which a comparative high advantage on a few variables can compensate a comparative disadvantage on many variables (Munda, 2008). We relaxed this strong perfect substitutability assumption by introducing a parametric family of aggregating functions that are known as generalized

weighted means. Parameterized by $\beta \in \mathfrak{R}$, the generalized weighted mean of a vector \mathbf{x} given weights \mathbf{w} is given by:

$$y^\beta(\mathbf{x}, \mathbf{w}) = \left(\sum_i w_i x_i^\beta \right)^{\frac{1}{\beta}}.$$

When $\beta = 1$ (0), the above function reduces to a weighted arithmetic (geometric) mean. The parameter β can be interpreted in terms of the elasticity of substitution between the different dimensions of the index, ϵ , where $\epsilon = \frac{1}{1-\beta}$. The smaller the value of β , the lower the substitutability between the different dimensions of performance (note that the case $\beta = 1$ corresponding to an arithmetic mean implies infinite substitutability).

For values of $\beta < 1$, generalized weighted means reflect a preference for balanced performance across the different dimensions of the index. Such balance is desirable in our context, so for the purposes of our uncertainty analysis we restricted ourselves to this range of β . Specifically, in our simulations we considered five values for β , namely $\beta \in \{0, 0.25, 0.50, .75, 1\}$, ranging from the arithmetic to the geometric mean.

3.3. Generating weight-aggregation samples

We generated a sample of 500 weight-aggregation pairs in the following manner. First, we draw a vector of objective weights $\mathbf{w} = (w_{EH}, w_{EV})$ where w_{EH} varied uniformly in $[0.3, 0.5]$ and $w_{EV} = 1 - w_{EH}$. Using these weights \mathbf{w} , country EPI scores are computed via their generalized weighted means for $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$, where the aggregations are performed at both the issue area and objective levels.⁵ For the aggregation of the issue areas we maintained the relative weights chosen by the developers of the 2014 EPI. For each value of β we estimated the Pearson correlation ratios for the two objectives with the techniques of Paruolo et al. (2013). If these correlation ratios were within 5% of each other for all β , then we kept this \mathbf{w} along with its corresponding five (\mathbf{w}, β) combinations. Thus, in

⁵ To avoid problems with the calculation of geometric means, we re-normalized issue area scores within the interval $[1, 100]$.

the end we obtained a sample of $100 \times 5 = 500$ weight-aggregation pairs where each sample ensures balanced importance for the two policy objectives with respect to the EPI.

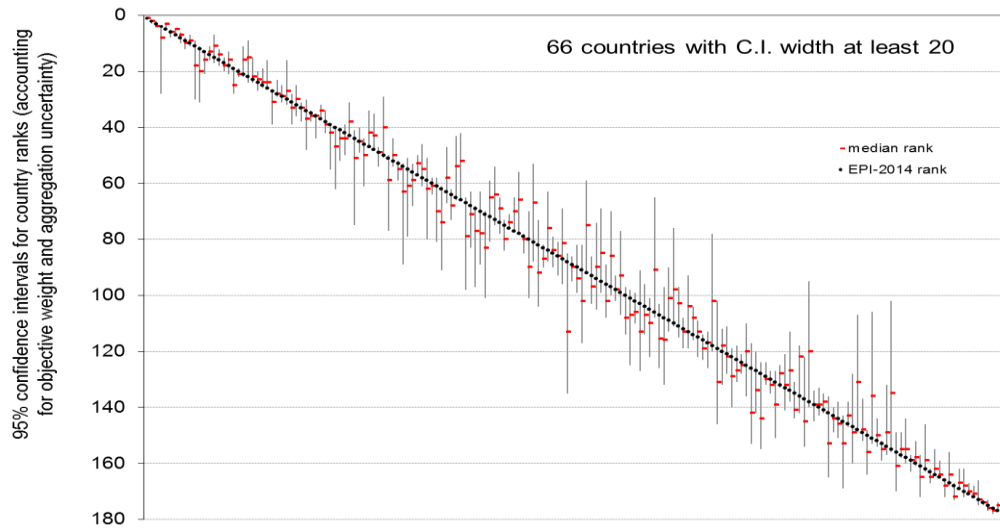
Table 4. Sources of uncertainty in the 2014 EPI: policy objective weights and aggregation function

	<i>Reference</i>	<i>Alternative</i>
I. Uncertainty in the aggregation formula	Weighted arithmetic average, i.e., $\beta = 1$	Generalized weighted mean $\beta \in \{0, 0.25, 0.50, 0.75, 1\}$
II. Uncertainty in the weights	Reference value for the weight	Distribution assigned for robustness analysis
Environmental Health	0.4	$w_{EH} \sim U[0.3, 0.5]$
Ecosystem Vitality	0.6	$w_{EV} = 1 - w_{EH}$

3.4. Uncertainty Analysis Results

Figure 3 below presents the results of our uncertainty analysis. Countries are ordered from best to worst according to their reference rank (black dot), the red dot being the median rank. All published 2014 EPI ranks lay within the simulated 95% confidence intervals. However, it is also true that country ranks vary significantly with changes in weights and aggregation function. Indeed, 38 countries have 95% confidence interval widths between 20 and 29. Confidence intervals widths for 20 countries lie between 30 and 39, for 6 countries between 40 and 49 (India, Solomon Islands, Pakistan, Swaziland, CAR, China, Vanuatu), and for 2 between 50 and 59 (Belize, India).

Figure 3: Uncertainty analysis results for 2014 EPI country ranks (based on 500 weight-aggregation pairs)



For full transparency and information, Table 5 reports the 2014 EPI country ranks together with the simulated median values and 95% confidence intervals in order to better appreciate the robustness of the results to the choice of weights and aggregation function. Confidence intervals wider than 20 are highlighted in red.

Table 5. Uncertainty analysis results for 2014 EPI country ranks

	EPI Rank	Median	95%CI		EPI Rank	Median	95%CI		EPI Rank	Median	95%CI
Switzerland	1	1	[1,1]	Seychelles	61	70	[58,81]	Zambia	121	122	[111,128]
Luxembourg	2	2	[2,3]	Montenegro	62	74	[61,91]	Papua New Guinea	122	129	[119,140]
Australia	3	4	[2,4]	Azerbaijan	63	58	[47,67]	Equatorial Guinea	123	127	[118,131]
Singapore	4	8	[4,28]	Cuba	64	68	[62,73]	Senegal	124	125	[121,128]
Czech Republic	5	3	[3,5]	Mexico	65	54	[43,66]	Kyrgyzstan	125	120	[110,136]
Germany	6	6	[5,8]	Turkey	66	52	[42,66]	Burkina Faso	126	142	[117,153]
Spain	7	5	[5,7]	Albania	67	79	[65,98]	Laos	127	134	[120,142]
Austria	8	7	[6,10]	Syria	68	71	[63,83]	Malawi	128	144	[124,155]
Sweden	9	10	[8,12]	Sri Lanka	69	77	[68,97]	Cote d'Ivoire	129	130	[124,132]
Norway	10	9	[7,11]	Uruguay	70	78	[63,89]	Congo	130	132	[127,135]
Netherlands	11	18	[9,30]	Suriname	71	83	[70,101]	Ethiopia	131	139	[127,151]
United Kingdom	12	20	[11,31]	South Africa	72	65	[59,81]	Timor-Leste	132	128	[125,133]
Denmark	13	16	[13,21]	Russia	73	64	[54,75]	Paraguay	133	132	[121,141]
Iceland	14	13	[11,16]	Moldova	74	69	[65,78]	Nigeria	134	127	[113,138]
Slovenia	15	11	[7,17]	Dominican Republic	75	80	[73,84]	Uganda	135	141	[131,144]
New Zealand	16	14	[12,18]	Fiji	76	74	[71,76]	Viet Nam	136	122	[118,142]
Portugal	17	18	[15,20]	Brazil	77	70	[65,77]	Guyana	137	145	[122,154]
Finland	18	16	[13,21]	Thailand	78	66	[56,78]	Swaziland	138	120	[95,140]
Ireland	19	25	[19,28]	Trinidad and Tobago	79	80	[78,85]	Nepal	139	139	[134,145]
Estonia	20	21	[19,22]	Palau	80	90	[70,101]	Kenya	140	139	[133,143]
Slovakia	21	16	[11,22]	Morocco	81	67	[53,88]	Cameroon	141	138	[135,142]
Italy	22	15	[9,24]	Bahrain	82	92	[73,104]	Niger	142	153	[136,165]
Greece	23	22	[15,24]	Iran	83	87	[82,90]	Tanzania	143	144	[140,149]
Canada	24	23	[20,27]	Kazakhstan	84	76	[63,88]	Guinea-Bissau	144	146	[138,151]
United Arab Emirates	25	24	[19,26]	Colombia	85	84	[79,90]	Cambodia	145	153	[143,169]
Japan	26	24	[16,26]	Romania	86	86	[81,93]	Rwanda	146	143	[138,147]
France	27	31	[27,39]	Bolivia	87	81.5	[69,96]	Grenada	147	149	[128,160]
Hungary	28	28	[23,29]	Belize	88	113	[85,135]	Pakistan	148	131	[107,149]
Chile	29	29	[25,31]	Macedonia	89	90	[86,95]	Iraq	149	148	[137,152]
Poland	30	27	[16,32]	Nicaragua	90	94	[82,99]	Benin	150	156	[148,164]
Serbia	31	33	[28,39]	Lebanon	91	102	[82,117]	Ghana	151	136	[106,153]
Belarus	32	30	[25,36]	Algeria	92	75	[59,92]	Solomon Islands	152	150	[144,154]
United States of America	33	33	[31,38]	Argentina	93	97	[86,103]	Comoros	153	155	[152,159]
Malta	34	37	[30,48]	Zimbabwe	94	90	[74,105]	Tajikistan	154	149	[132,157]
Saudi Arabia	35	36	[34,38]	Ukraine	95	85	[69,99]	India	155	135	[102,156]
Belgium	36	36	[35,44]	Antigua and Barbuda	96	102	[89,108]	Chad	156	161	[149,170]
Brunei Darussalam	37	34	[32,37]	Honduras	97	86	[70,100]	Yemen	157	155	[149,158]
Cyprus	38	39	[34,42]	Guatemala	98	98	[93,102]	Mozambique	158	155	[144,160]
Israel	39	42	[39,55]	Oman	99	93	[77,107]	Gambia	159	159	[157,163]
Latvia	40	47	[40,62]	Botswana	100	108	[97,114]	Angola	160	158	[152,161]
Bulgaria	41	44	[40,52]	Georgia	101	107	[98,125]	Djibouti	161	165	[157,172]
Kuwait	42	44	[39,50]	Dominica	102	106	[99,110]	Guinea	162	159	[146,163]
South Korea	43	38	[31,43]	Bhutan	103	113	[91,127]	Togo	163	165	[162,167]
Qatar	44	51	[41,75]	Gabon	104	107	[96,114]	Myanmar	164	162	[155,164]
Croatia	45	45	[40,49]	Bahamas	105	110	[101,122]	Mauritania	165	164	[159,165]
Taiwan	46	50	[44,61]	Vanuatu	106	91	[65,108]	Madagascar	166	168	[166,172]
Tonga	47	42	[34,47]	Bosnia and Herzegovina	107	115.5	[103,126]	Burundi	167	164	[156,168]
Armenia	48	43	[35,48]	Barbados	108	116	[97,132]	Eritrea	168	172	[167,173]
Lithuania	49	49	[47,54]	Turkmenistan	109	101	[90,113]	Bangladesh	169	167	[162,169]
Egypt	50	40	[29,51]	Peru	110	98	[76,111]	Dem. Rep. Congo	170	168	[162,172]
Malaysia	51	59	[50,77]	Mongolia	111	103	[97,115]	Sudan	171	170	[167,171]
Tunisia	52	50	[44,54]	Indonesia	112	113	[108,119]	Liberia	172	171	[168,172]
Ecuador	53	55	[52,59]	Cape Verde	113	104	[93,119]	Sierra Leone	173	173	[166,175]
Costa Rica	54	63	[53,89]	Philippines	114	108	[104,116]	Afghanistan	174	174	[173,175]
Jamaica	55	61	[51,79]	El Salvador	115	113	[109,122]	Lesotho	175	176	[173,177]
Mauritius	56	59	[53,68]	Namibia	116	119	[114,123]	Haiti	176	177	[176,178]
Venezuela	57	53	[50,57]	Uzbekistan	117	117	[113,126]	Mali	177	175	[174,177]
Panama	58	55	[46,61]	China	118	102	[78,120]	Somalia	178	178	[176,178]
Kiribati	59	62	[51,80]	Central African Republic	119	131	[102,146]				
Jordan	60	61	[58,64]	Libya	120	118	[112,132]				

3.5. The relative importance of uncertain input factors to the variation in 2014 EPI ranks

In this section we will investigate the relative importance of uncertainty in weights and aggregation in the 2014 EPI. As the following analysis will make clear, variation in country ranks is overwhelmingly driven by the choice of aggregation function.

Following Saisana et al. (2005), our measure of robustness is the absolute shift in rank with respect to the benchmark choice of equal weights and linear aggregation, which we denote by the variable ΔR . That is, given a country c and a weight-aggregation pair (w, β) , we are interested in the following quantity (here, $Rank_c(w, \beta)$ denotes country c 's rank under the version of our composite index that uses weights w and aggregation β):

$$\Delta R_c(w, \beta) = |Rank_c(w^{ref}, 1) - Rank_c(w, \beta)|.$$

Given a weight-aggregation pair (w, β) , a compelling aggregate measure of robustness can be found in the *average* shift in rank (over the set of countries) that (w, β) results in, denoted by $\mu_{\Delta R}(w, \beta)$, (here C is the number of countries):

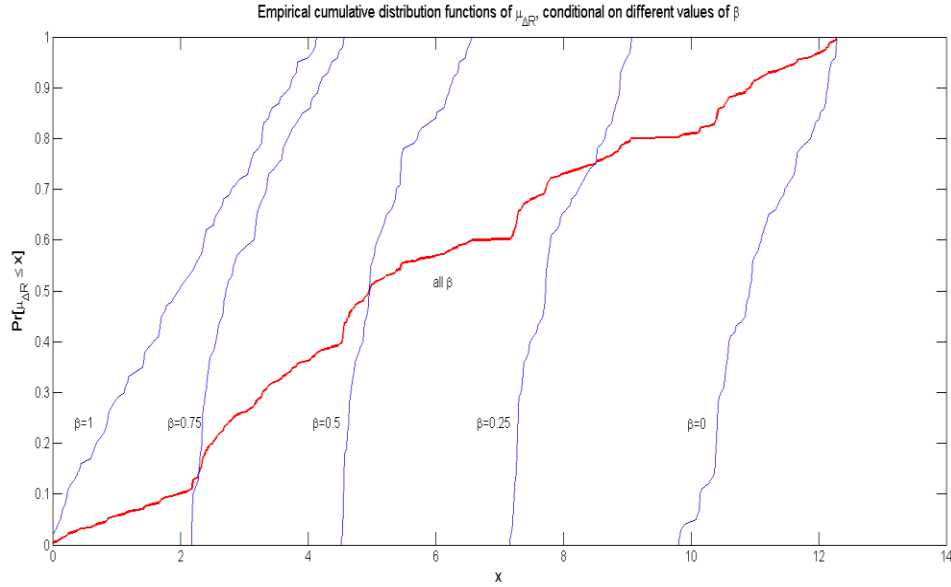
$$\mu_{\Delta R}(w, \beta) = \frac{1}{C} * \sum_{c=1}^C \Delta R_c(w, \beta)$$

Focusing on our simulated sample, the sample mean and standard deviation for $\mu_{\Delta R}$ are given by: $\overline{\mu_{\Delta R}} = 5.8$, $\sigma_{\mu_{\Delta R}} = 3.4$. Zooming in now on the choice of aggregation, we denote by $\overline{\mu_{\Delta R}}(\beta)$ and $\sigma_{\mu_{\Delta R}}(\beta)$ the expectation and sample standard deviation of $\mu_{\Delta R}$ conditional on different values of β . We have:

β	$\overline{\mu_{\Delta R}}(\beta)$	$\sigma_{\mu_{\Delta R}}(\beta)$
1	2.0	1.2
0.75	3.0	0.7
0.50	5.2	0.6
0.25	7.9	0.6
0	11.0	0.7

Figure 4 below depicts the empirical cumulative distribution function (cdf) of $\mu_{\Delta R}$, as well as the analogous distributions conditional on the 5 values of β :

Figure 4: Empirical cumulative distribution function of mean shift in rank.



Note: This figure can be read in the following way. Suppose we are interested in the p^{th} percentiles of the conditional and unconditional distributions of $\mu_{\Delta R}$, where the conditioning is performed on the choice of aggregation function. Then, draw a straight horizontal line originating at point p on the y-axis. This line will intersect the 5 conditional (blue) and 1 unconditional (red) cdfs at different points, and the x-coordinates of these points will be the p^{th} percentiles of the respective distributions. For instance, conditional on $\beta = 0.25$, 75 percent of the simulated EPI rankings have an average absolute shift in rank of at most 8.5, with respect to the original EPI rankings. By chance, this value happens to also be the 75th percentile of the unconditional distribution.

Figure 4 makes graphically clear how the choice of aggregation is the main driver behind the variation in country ranks. If we fix a value for β , we see that the resulting cdfs are very concentrated, with a support length of about 2, with the exception of $\beta = 1$ for which it is around 4. Contrast these results with the unconditional cdf shown in red, whose support length is greater than 10.

This point can be made also algebraically. Define the *sensitivity index* $S_w(S_\beta)$ to be the fractional contribution to the sample variance of $\mu_{\Delta R}$ due to the uncertainty in the weights (aggregation scheme) of the EPI. Equivalently, let $S_{w\beta}$ denote this contribution due to the

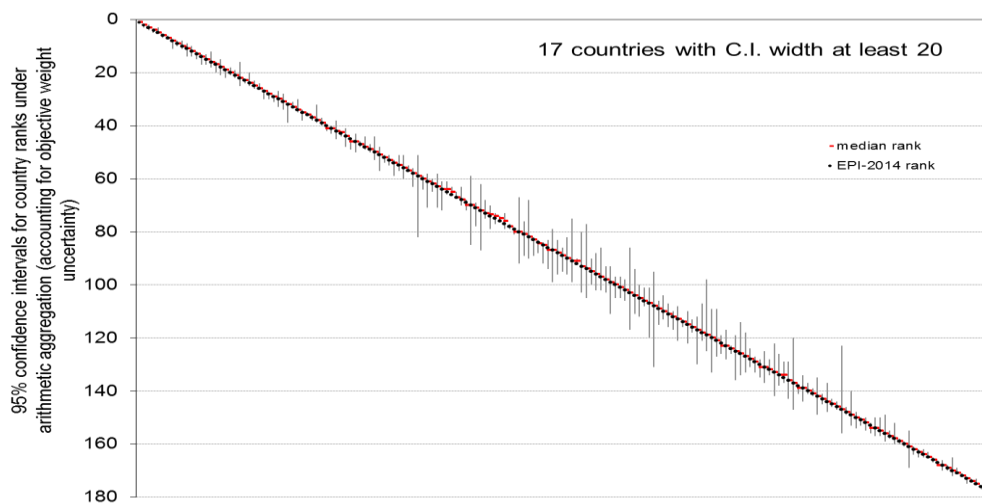
interaction effect of uncertainty in both weights and aggregation.⁶ Simple calculations yield: $S_w = 0.04$, $S_\beta = 0.94$, $S_{w\beta} = 0.02$.

Thus, we see that the choice of aggregation function is responsible for 94% of the sample variance of $\mu_{\Delta R}$, while the choice of weights only for 4%. Indeed, country ranks seem to be robust to the choice of policy objective weights, but quite sensitive to the choice of aggregation function.

3.6. Uncertainty analysis under fixed arithmetic aggregation

Given the above results, we may be interested in asking how robust are the 2014 EPI ranks under exclusively arithmetic aggregation. In Figure 4, the conditional distribution of $\mu_{\Delta R}$ given $\beta = 1$ suggests a significant degree of robustness, but we wish to investigate this point a little further. Figure 5 shows the simulated country ranks given a fixed choice of arithmetic aggregation. Indeed, comparing it to Figure 3, we see that confidence intervals are much narrower, with only 17 countries having a width of 20 or above. Out of those, 13 countries are in the 20-29 range, and 4 in the 30-39 (Kiribati, Bhutan, Barbados, and Grenada).

Figure 5: Uncertainty analysis results under fixed arithmetic aggregation.



⁶ For details on the precise definition of sensitivity indices see Saisana et al. (2005).

The primary factor behind the wide confidence intervals for these countries is uneven performance in the two policy objectives. For instance, in the original 2014 EPI ranking, Kiribati's rank for Environmental Health is 122, while for Ecosystem Vitality is 28, and the corresponding figures for Bhutan, Barbados, and Grenada are (136,44), (39,166), and (68,175), respectively. These are very big swings. Therefore, even small differences in the weights of the two objectives, combined with the perfect substitutability of arithmetic aggregation, will result in an arithmetic average that is substantially different from that of the 2014 EPI ranking.

For completeness, Table 6 below presents the uncertainty analysis results for each country for the entire sample, as well as the restricted sample corresponding to fixed arithmetic means. Once again, confidence intervals greater than 20 are highlighted in red. The higher robustness of the $\beta = 1$ case is apparent. However, it is also worth noting that the confidence intervals of 9 countries (Kiribati, South Africa, Argentina, Guatemala, Barbados, Uzbekistan, Libya, Zambia, and Grenada) become wider under fixed arithmetic aggregation. This suggests that weight uncertainty does play a role in the observed variance of the 2014 EPI ranks.

Table 6: 2014 EPI: Uncertainty analysis results with/without fixed arithmetic aggregation

	all β					$\beta=1$						all β					$\beta=1$				
	EPI Rank	Median	95%CI	Median	95% CI		EPI Rank	Median	95%CI	Median	95% CI		EPI Rank	Median	95%CI	Median	95% CI				
Switzerland	1	1	[1,1]	1	[1,1]	Seychelles	61	70	[58,81]	61	[58,71]	Zambia	121	122	[111,128]	121	[109,127]				
Luxembourg	2	2	[2,3]	2	[2,3]	Montenegro	62	74	[61,91]	62	[60,65]	Papua New Guinea	122	129	[119,140]	123	[119,126]				
Australia	3	4	[2,4]	3	[2,4]	Azerbaijan	63	58	[47,67]	63	[58,71]	Equatorial Guinea	123	127	[118,131]	123	[117,128]				
Singapore	4	8	[4,28]	4	[4,5]	Cuba	64	68	[62,73]	64	[61,72]	Senegal	124	125	[121,128]	124	[123,126]				
Czech Republic	5	3	[3,5]	5	[3,5]	Mexico	65	54	[43,66]	64	[61,66]	Kyrgyzstan	125	120	[110,136]	125	[119,136]				
Germany	6	6	[5,8]	6	[6,7]	Turkey	66	52	[42,66]	65	[63,66]	Burkina Faso	126	142	[117,153]	126	[114,134]				
Spain	7	5	[5,7]	7	[6,7]	Albania	67	79	[65,98]	67	[65,67]	Laos	127	134	[120,142]	127	[118,133]				
Austria	8	7	[6,10]	8	[8,11]	Syria	68	71	[63,83]	68	[63,70]	Malawi	128	144	[124,155]	128	[124,131]				
Sweden	9	10	[8,12]	9	[8,10]	Sri Lanka	69	77	[68,97]	70	[68,72]	Cote d'Ivoire	129	130	[124,132]	129	[127,133]				
Norway	10	9	[7,11]	10	[8,11]	Uruguay	70	78	[63,89]	70	[59,85]	Congo	130	132	[127,135]	131	[128,135]				
Netherlands	11	18	[9,30]	11	[9,14]	Suriname	71	83	[70,101]	71	[69,78]	Ethiopia	131	139	[127,151]	131	[125,137]				
United Kingdom	12	20	[11,31]	12	[10,14]	South Africa	72	65	[59,81]	72	[62,87]	Timor-Leste	132	128	[125,133]	132	[128,134]				
Denmark	13	16	[13,21]	13	[12,15]	Russia	73	64	[54,75]	73	[68,75]	Paraguay	133	132	[121,141]	133	[122,142]				
Iceland	14	13	[11,16]	14	[13,17]	Moldova	74	69	[65,78]	73.5	[70,79]	Nigeria	134	127	[113,138]	134	[126,138]				
Slovenia	15	11	[7,17]	15	[13,17]	Dominican Republic	75	80	[73,84]	74	[73,77]	Uganda	135	141	[131,144]	134	[129,137]				
New Zealand	16	14	[12,18]	16	[12,18]	Fiji	76	74	[71,76]	75	[74,76]	Viet Nam	136	122	[118,142]	136	[129,143]				
Portugal	17	18	[15,20]	17	[16,20]	Brazil	77	70	[65,77]	76	[73,79]	Guyana	137	145	[122,154]	137	[120,147]				
Finland	18	16	[13,21]	18	[15,21]	Thailand	78	66	[56,78]	78	[77,79]	Swaziland	138	120	[95,140]	139	[138,141]				
Ireland	19	25	[19,28]	19	[19,22]	Trinidad and Tobago	79	80	[78,85]	80	[79,81]	Nepal	139	139	[134,145]	139	[134,144]				
Estonia	20	21	[19,22]	20	[18,21]	Palau	80	90	[70,101]	80	[67,92]	Kenya	140	139	[133,143]	140	[137,141]				
Slovakia	21	16	[11,22]	21	[19,22]	Morocco	81	67	[53,88]	81	[76,89]	Cameroon	141	138	[135,142]	141	[139,142]				
Italy	22	15	[9,24]	22	[16,25]	Bahrain	82	92	[73,104]	82	[68,90]	Niger	142	153	[136,165]	142	[135,149]				
Greece	23	22	[15,24]	23	[23,24]	Iran	83	87	[82,90]	83	[82,89]	Tanzania	143	144	[140,149]	143	[141,145]				
Canada	24	23	[20,27]	24	[20,25]	Kazakhstan	84	76	[63,88]	84.5	[83,88]	Guinea-Bissau	144	146	[138,151]	144	[137,148]				
United Arab Emirates	25	24	[19,26]	25	[23,26]	Colombia	85	84	[79,90]	85	[84,92]	Cambodia	145	153	[143,169]	145	[143,146]				
Japan	26	24	[16,26]	26	[24,26]	Romania	86	86	[81,93]	87	[83,94]	Rwanda	146	143	[138,147]	146	[144,147]				
France	27	31	[27,39]	27	[27,30]	Bolivia	87	81.5	[69,96]	87	[79,99]	Grenada	147	149	[128,160]	147	[123,156]				
Hungary	28	28	[23,29]	28	[28,30]	Belize	88	113	[85,135]	88	[83,96]	Pakistan	148	131	[107,149]	148	[146,149]				
Chile	29	29	[25,31]	29	[29,31]	Macedonia	89	90	[86,95]	89	[86,95]	Iraq	149	148	[137,152]	149	[140,153]				
Poland	30	27	[16,32]	30	[27,33]	Nicaragua	90	94	[82,99]	90	[82,96]	Benin	150	156	[148,164]	150	[148,154]				
Serbia	31	33	[28,39]	31	[28,34]	Lebanon	91	102	[82,117]	91	[75,99]	Ghana	151	136	[106,153]	151	[150,153]				
Belarus	32	30	[25,36]	32	[32,39]	Algeria	92	75	[59,92]	91	[90,92]	Solomon Islands	152	150	[144,154]	152	[150,155]				
United States of America	33	33	[31,38]	33	[31,33]	Argentina	93	97	[86,103]	93	[80,103]	Comoros	153	155	[152,159]	154	[152,156]				
Malta	34	37	[30,48]	34	[30,35]	Zimbabwe	94	90	[74,105]	94	[77,105]	Tajikistan	154	149	[132,157]	154	[150,157]				
Saudi Arabia	35	36	[34,38]	35	[34,38]	Ukraine	95	85	[69,99]	95	[90,100]	India	155	135	[102,156]	155	[150,157]				
Belgium	36	36	[35,44]	36	[35,36]	Antigua and Barbuda	96	102	[89,108]	96	[88,102]	Chad	156	161	[149,170]	156	[149,159]				
Brunei Darussalam	37	34	[32,37]	37	[36,38]	Honduras	97	86	[70,100]	97	[86,102]	Yemen	157	155	[149,158]	157	[155,158]				
Cyprus	38	39	[34,42]	38	[32,39]	Guatemala	98	98	[93,102]	98	[93,103]	Mozambique	158	155	[144,160]	158	[152,160]				
Israel	39	42	[39,55]	39	[37,40]	Oman	99	93	[77,107]	99	[93,111]	Gambia	159	159	[157,163]	159	[158,160]				
Latvia	40	47	[40,62]	41	[40,42]	Botswana	100	108	[97,114]	100	[97,105]	Angola	160	158	[152,161]	160	[159,161]				
Bulgaria	41	44	[40,52]	41	[40,43]	Georgia	101	107	[98,125]	101	[97,105]	Djibouti	161	165	[157,172]	161	[155,169]				
Kuwait	42	44	[39,50]	42	[38,45]	Dominica	102	106	[99,110]	102	[98,106]	Guinea	162	159	[146,163]	162	[161,164]				
South Korea	43	38	[31,43]	42.5	[42,44]	Bhutan	103	113	[91,127]	103	[86,117]	Togo	163	165	[162,167]	163	[162,165]				
Qatar	44	51	[41,75]	44	[41,48]	Gabon	104	107	[96,114]	104	[94,111]	Myanmar	164	162	[155,164]	164	[162,165]				
Croatia	45	45	[40,49]	46	[45,49]	Bahamas	105	110	[101,122]	105	[100,112]	Mauritania	165	164	[159,165]	165	[163,165]				
Taiwan	46	50	[44,61]	46	[43,50]	Vanuatu	106	91	[65,108]	106	[101,108]	Madagascar	166	168	[166,172]	166	[165,167]				
Tonga	47	42	[34,47]	47	[46,48]	Bosnia and Herzegovina	107	115.5	[103,126]	107	[101,120]	Burundi	167	164	[156,168]	168	[167,168]				
Armenia	48	43	[35,48]	48	[44,49]	Barbados	108	116	[97,132]	108	[95,131]	Eritrea	168	172	[167,173]	168	[166,170]				
Lithuania	49	49	[47,54]	49	[47,50]	Turkmenistan	109	101	[90,113]	109	[106,115]	Bangladesh	169	167	[162,169]	169	[168,170]				
Egypt	50	40	[29,51]	50	[44,53]	Peru	110	98	[76,111]	110	[104,113]	Dem. Rep. Congo	170	168	[162,172]	170	[165,172]				
Malaysia	51	59	[50,77]	51	[48,57]	Mongolia	111	103	[97,115]	111	[107,116]	Sudan	171	170	[167,171]	171	[169,171]				
Tunisia	52	50	[44,54]	52	[51,54]	Indonesia	112	113	[108,119]	112	[110,117]	Liberia	172	171	[168,172]	172	[171,172]				
Ecuador	53	55	[52,59]	53	[51,55]	Cape Verde	113	104	[93,119]	113	[108,121]	Sierra Leone	173	173	[166,175]	173	[173,175]				
Costa Rica	54	63	[53,89]	54	[53,59]	Philippines	114	108	[104,116]	114	[113,116]	Afghanistan	174	174	[173,175]	174	[173,174]				
Jamaica	55	61	[51,79]	55	[51,57]	El Salvador	115	113	[109,122]	115	[110,122]	Lesotho	175	176	[173,177]	175	[173,175]				
Mauritius	56	59	[53,68]	56	[51,60]	Namibia	116	119	[114,123]	116	[113,118]	Haiti	176	177	[176,178]	176	[176,177]				
Venezuela	57	53	[50,57]	57	[55,57]	Uzbekistan	117	117	[113,126]	117	[112,130]	Mali	177	175	[174,177]	177	[176,177]				
Panama	58	55	[46,61]	58	[56,63]	China	118	102	[78,120]	118	[107,121]	Somalia	178	178	[176,178]	178	[178,178]				
Kiribati	59	62	[51,80]	59	[51,82]	Central African Republic	119	131	[102,146]	119	[98,125]										
Jordan	60	61	[58,64]	60	[59,64]	Libya	120	118	[112,132]	120	[109,133]										

4. Conclusions

The 2014 EPI developing team solicited the input of the JRC to investigate the statistical properties of the index in order to ensure the transparency and reliability of the index and enable policymakers to derive more accurate and meaningful conclusions. To this end, we checked (1) the statistical coherence of the EPI and (2) its robustness to modeling assumptions regarding the weight and aggregation functions. Our analysis enabled us to formulate the following recommendations and conclusions.

First, our analysis showed that the 2014 EPI is well-balanced with respect to the two policy objectives and that these objectives are also adequately correlated to justify their aggregation into an overall index. This is a noteworthy improvement compared to past releases. Second, we observed good to strong correlations between indicators and the respective EPI issue areas, which implies meaningful indicator contribution to the variance of the aggregate scores. Third, the indicators' correlation structure within and across the nine EPI issue areas suggests that the indicators have been allocated to the most relevant environmental issue.

Nevertheless, regardless of the conceptual relevance to the phenomenon being measured, three EPI issue areas – Forests, Fisheries and Agriculture – need a careful revision, possibly by populating them with different indicators. As they stand now, they do not meaningfully contribute to the variation in the Ecosystem Vitality scores or in the EPI scores, and instead seem to add noise to their measurement. Furthermore, since we believe that the 2014 EPI should be composed of indicators that “have an impact” in the classification of country performance, we recommend a careful revision of the following indicators: (a) average Exposure to PM_{2.5} and PM_{2.5} Exceedance (Air Quality), (b) Change of Trend in Carbon Intensity (Climate and Energy), and (c) Agricultural Subsidies (Agriculture).

We understand that, due to data quality and availability limitations, these recommendations cannot be easily adopted. As mentioned in Hsu et al. (2014), the EPI developers are aware of the data issues, including ones identified by the JRC as problematic, such as those related to the Agriculture issue area. The developing team tried to reduce the impact of these indicators on the EPI scores by, for example, granting them lower weights, see Hsu et al. (2014, p. 20). Our analysis suggests that there is more work to be done on this front.

Second, examining the main effects (sensitivity indices) for the two EPI policy objectives, we found them to be strong and almost perfectly balanced. This result suggests that the weighting scheme chosen by the developing team has indeed led to the desired outcome on the importance of the two policy objectives in classifying countries in the 2014 EPI. At the issue area level, however, the 2014 EPI appears to be less balanced. Within the Environmental Health objective, the main effect for Air Quality is significantly smaller than that of Health Impacts and Water & Sanitation, despite their equal weights within the objective. Conversely, within the Ecosystem Vitality objective, we again notice an uneven pattern. For instance, Agriculture has a main effect of 0.27, while Forests and Fisheries are practically insignificant, despite the fact that they are weighted twice as much. Similarly, Water Resources and Climate & Energy have significantly different values, despite their equal weight. Moving to the issue areas importance measures for the 2014 EPI as a whole, the negative results of the within-objective analysis remain. Indeed, they are exacerbated.

Thirdly, we investigated the robustness of 2014 EPI country ranks to the choice of policy objectives weights and aggregation function via an uncertainty and sensitivity analysis. In doing so, we simulated 500 weight-aggregation pairs that ensured balanced policy objective contributions. We found that all published EPI ranks lay within the simulated 95% confidence intervals. Nevertheless, we also noticed that country ranks vary significantly with changes in weights and aggregation function: 38 countries have 95% confidence interval widths between 20 and 29, 20 countries between 30 and 39, 6 countries between 40 and 49 (India, Solomon Islands, Pakistan, Swaziland, CAR, China, Vanuatu), and 2 between 50 and 59 (Belize, India). For those countries, the 2014 EPI ranks need to be treated with caution.

The choice of aggregation function is the main driver behind this variation in country ranks. Choosing the average absolute shift in rank as our robustness metric, we saw that the choice of aggregation function is responsible for 94% of the sample variance, while the choice of weights only for 4%. This result suggests that should the methodological choices behind the 2014 EPI stimulate further discussions, then these should focus more on the aggregation formula for the two policy objectives and much less on their weights. However, it is also worth noting that the confidence intervals of 9 countries (Kiribati, South Africa, Argentina, Guatemala, Barbados, Uzbekistan, Libya, Zambia, and Grenada) become wider under fixed

arithmetic aggregation. This suggests that weight uncertainty does play a role in the observed variance of the 2014 EPI ranks for specific countries, though not for the majority of them. The primary factor behind the sensitivity in these countries' ranks is uneven performance in the two policy objectives. For instance, in the original 2014 EPI ranking, Kiribati's rank for Environmental Health is 122, while for Ecosystem Vitality is 28, and the corresponding figures for Bhutan, Barbados, and Grenada are (136,44), (39,166), and (68,175), respectively. These are very big swings. Therefore, even small differences in the weights of the two EPI objectives, combined with the perfect substitutability of arithmetic aggregation, will result in an arithmetic average that is substantially different from that of the 2014 EPI ranking.

The auditing conducted herein has shown the potential of the 2014 Environmental Performance Index, upon some further refinements, in reliably identifying weaknesses and ultimately monitoring national performance in high-priority environmental issues around the world.

References

1. Decancq, K., & Lugo, M. A. 2013. Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*, 32(1), 7-34.
2. Groeneveld, R.A. and Meeden, G. 1984. Measuring skewness and kurtosis. *The Statistician* 33, 391-399.
3. Hsu, A., J. Emerson, M. Levy, A. de Sherbinin, L. Johnson, O. Malik, J. Schwartz, and M. Jaiteh. 2014. The 2014 Environmental Performance Index. New Haven, CT: Yale Center for Environmental Law & Policy. Available: www.epi.yale.edu.
4. Little, R.J.A., Rubin, D.B. 2002. Statistical Analysis with missing data. 2nd edition; John Wiley & Sons, Inc.
5. Munda, G. 2008. Social Multi-Criteria Evaluation for a Sustainable Economy. Berlin Heidelberg: Springer-Verlag.
6. OECD/EC JRC. 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. Paris: OECD.

7. Paruolo P., Saisana M., Saltelli A. 2013. Ratings and Rankings: voodoo or science?. *J Royal Statistical Society A* 176(3), 609-634.
8. Saisana, M., and Saltelli, A. 2011. Rankings and Ratings: Instructions for use, *Hague Journal on the Rule of Law* 3(2), 247-268.
9. Saisana, M., D'Hombres, B., Saltelli, A. 2011. Rickety numbers: Volatility of university rankings and policy implications, *Research Policy* 40, 165-177.
10. Saisana, M., Saltelli, A., Tarantola, S. 2005. Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators. *Journal of the Royal Statistical Society A* 168(2), 307-323.
11. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. 2008. *Global Sensitivity Analysis: The Primer*. Chichester, England: John Wiley & Sons.

Table A-1. Pearson's correlation coefficients – 2014 EPI indicators

	Category	Indicator	(1)	(2)	(3)	(4)	(5)
Environment Health	Health Impacts	Child Mortality (1)	1.00				
	Air Quality	Household Air Quality (2)	0.82	1.00			
		Air Pollution - Average Exposure to PM2.5 (3)	-0.20	-0.06 ^{ns}	1.00		
		Air Pollution - PM2.5 Exceedance (4)	-0.28	-0.15 ^{ns}	0.95	1.00	
	Water & Sanitation	Access to Drinking Water (5)	0.86	0.75	-0.19	-0.27	1.00
		Access to Sanitation (6)	0.80	0.77	-0.13	-0.24	0.81

	Category	Indicator	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Ecosystem Vitality	Water Resources	Wastewater Treatment (7)	1.00											
	Agriculture	Agricultural Subsidies (8)	-0.58	1.00										
		Pesticide Regulation (9)	0.33	-0.41*	1.00									
	Forests	Change in Forest Cover (10)	0.12	0.00	-0.05	1.00								
	Fisheries	Coastal Shelf Fishing Pressure (11)	-0.19	0.14	-0.21	-0.02	1.00							
		Fish Stocks (12)	-0.08	-0.18	0.06	-0.06	0.28	1.00						
	Biodiversity & Habitat	Terrestrial Protected Areas (National Biome Weights) (13)	0.14	-0.29	0.16	-0.34	-0.02	0.12	1.00					
		Terrestrial Protected Areas (Global Biome Weights) (14)	0.14	-0.30	0.16	-0.32	-0.04	0.08	0.97	1.00				
		Marine Protected Areas (15)	0.25	-0.40	0.28	-0.13	0.02	0.27	0.36	0.37	1.00			
		Critical Habitat Protection (16)	-0.03	0.02	0.24	-0.10	0.09	0.11	0.47	0.50	0.26	1.00		
	Climate & Energy	Trend in Carbon Intensity (17)	0.18	-0.14	0.17	-0.01	-0.26	0.13	0.06	0.10	0.19	0.02	1.00	
		Change of Trend in Carbon Intensity (18)	-0.15	0.13	-0.05	0.04	-0.08	0.18	-0.06	-0.07	0.06	-0.04	0.15 ^{ns}	1.00
		Trend in CO2 Emissions per KWH (19)	0.30	-0.15	0.16	-0.03	0.01	-0.14	-0.11	-0.08	-0.06	0.18	0.29	-0.05 ^{ns}

Note: * indicates undesirable negative correlation; 'ns' indicates not significant correlation at 99% level

Table A-2. Correlations between 2014 EPI components under the assumption of a geometric averaging

Policy objective	Policy category	Indicator	correlation between indicator and objective	correlation between indicator and EPI	correlation between issue area and objective	correlation between objective and EPI
Environmental Health	Health Impacts	Child Mortality	0.93	0.76	0.93	0.77
	Air Quality	Household Air Quality	0.88	0.69	0.56	
		Air Pollution - Average Exposure to PM2.5	-0.03 ^{ns}	-0.07 ^{ns}		
		Air Pollution - PM2.5 Exceedance	-0.12 ^{ns}	-0.14 ^{ns}		
	Water & Sanitation	Access to Drinking Water	0.92	0.73	0.96	
		Access to Sanitation	0.90	0.70		
Ecosystem Vitality	Water Resources	Wastewater Treatment	0.82	0.83	0.82	0.98
	Agriculture	Agricultural Subsidies	-0.57*	-0.60*	-0.05 ^{ns}	
		Pesticide Regulation	0.41	0.45		
	Forests	Change in Forest Cover	0.09 ^{ns}	0.12 ^{ns}	0.22	
	Fisheries	Coastal Shelf Fishing Pressure	-0.05 ^{ns}	-0.05 ^{ns}	0.10 ^{ns}	
		Fish Stocks	0.06 ^{ns}	0.07 ^{ns}		
	Biodiversity & Habitat	National Biome Weights	0.35	0.33	0.42	
		Global Biome Weights	0.36	0.34		
		Marine Protected Areas	0.47	0.47		
		Critical Habitat Protection	0.17 ^{ns}	0.20 ^{ns}		
	Climate & Energy	Trend in Carbon Intensity	0.19 ^{ns}	0.17 ^{ns}	0.48	
		Change of Trend in Carbon Intensity	-0.07 ^{ns}	-0.10 ^{ns}		
		Trend in CO2 Emissions per KWH	0.23	0.26		

Note: * indicates undesirable negative correlation; 'ns' indicates not significant correlation at 99% level

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*): 00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.

It can be accessed through the Europa server <http://europa.eu/>.

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu/>), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents.

You can obtain their contact details by sending a fax to (352) 29 29-42758.

European Commission

EUR 26623 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Environmental Performance Index 2014 - JRC Analysis and Recommendations

Author(s): Stergios Athanasoglou, Dorota Weziak-Bialowolska, Michaela Saisana

Luxembourg: Publications Office of the European Union

2014 – 37 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1831-9424 (online)

ISBN 978-92-79-37915-4 (PDF)

doi: 10.2788/64170

Abstract

The latest edition of the Environmental Performance Index (EPI) was presented and discussed on January 2014 during the World Economic Forum (WEF) Annual Meeting in Davos. The EPI is released biannually since 2006 by Yale and Columbia Universities, in collaboration with the Samuel Foundation and the WEF. The EPI ranks how well countries perform on high-priority environmental issues concerning the policy areas of environmental health and ecosystem vitality.

The JRC's Econometrics and Applied Statistics Unit was invited for a fifth consecutive time to perform a statistical audit the EPI, focusing on two main questions:

1) Is the EPI multi-level structure statistically coherent?

2) What is the impact of modelling assumptions on the 2014 EPI ranking?

The 2014 EPI was found to be well-balanced with respect to its two policy objectives, which were also adequately correlated to justify their aggregation into an overall index. Satisfactory correlations were observed between indicators and respective EPI issue areas, implying meaningful indicator contributions to the variance of the aggregate scores. Possible refinements of the index mainly concern the issue areas of Forests, Fisheries and Agriculture, which do not seem to contribute significantly to the EPI ranking.

The JRC's uncertainty analysis investigated the robustness of EPI country ranks to two key choices: policy objective weights and aggregation function. The choice of aggregation function at the policy objectives level was found to be the main driver of the variation in country ranks, accounting for a much greater share of the observed variance in country ranks. This suggested that future deliberations on the index's methodology should focus primarily on the choice of aggregation function.

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

*Serving society
Stimulating innovation
Supporting legislation*